

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
21. Oktober 2004 (21.10.2004)

PCT

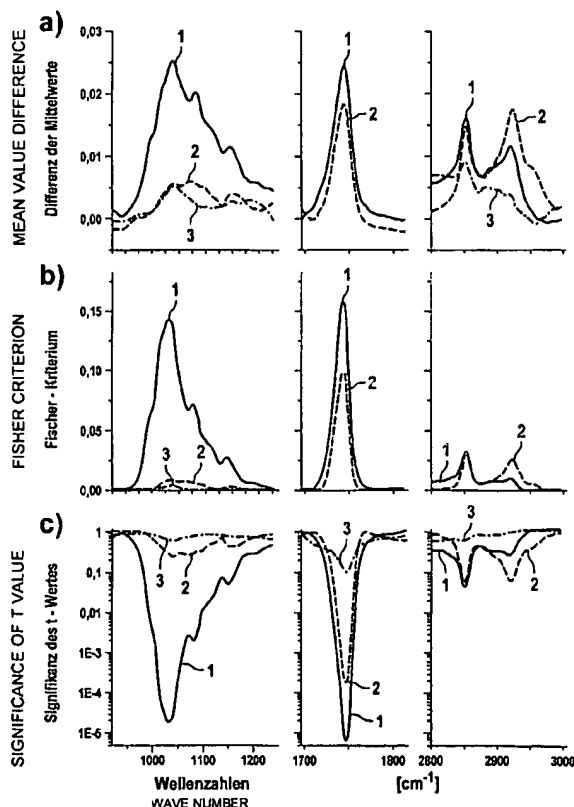
(10) Internationale Veröffentlichungsnummer
WO 2004/090779 A2

- (51) Internationale Patentklassifikation⁷: **G06F 19/00** (71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von DE): **F. HOFFMANN-LA-ROCHE AG** [CH/CH]; Gren-
(21) Internationales Aktenzeichen: **PCT/EP2004/003613** zacherstrasse 124, CH-4070 Basel (CH).
(22) Internationales Anmeldedatum: 6. April 2004 (06.04.2004) (72) Erfinder; und
(25) Einreichungssprache: Deutsch (75) Erfinder/Anmelder (nur für US): **PETRICH, Wolfgang**
(26) Veröffentlichungssprache: Deutsch [DE/DE]; Gutenbergstrasse 7a, 76669 Bad Schoenborn
(30) Angaben zur Priorität: 103 15 877.4 8. April 2003 (08.04.2003) DE (DE). **WERNER, Gerhard** [DE/DE]; Nachstenbacher
Weg 5, 69469 Weinheim (DE). **HEIN, Heinz-Michael**
(71) Anmelder (nur für DE): **ROCHE DIAGNOSTICS** [DE/DE]; Sandhofer Str. 116, 68305 Mannheim (DE). **KUHR, Hans-Juergen** [DE/DE]; Perreystrasse 30, 68219
Mannheim (DE).
(74) Gemeinsamer Vertreter: **WILDSCHÜTZ, Sabine**;
Roche Diagnostics GmbH, 68298 Mannheim (DE).

[Fortsetzung auf der nächsten Seite]

(54) Title: CONTROL FOR A DISEASE PROGRESSION

(54) Bezeichnung: KRANKHEITSVERLAUFKONTROLLE



(57) Abstract: The invention relates to analysis methods for diagnosing diseases on human and animal samples. Said invention also relates to an evaluation method for diagnosing the individual stages of a disease in such a way that it is possible to display the progression thereof. Said invention also makes it possible to identify diseases in an early manner and to carry out therapeutic controls. The inventive method consists in carrying out actually known multivariable evaluation methods for classifying samples. Nevertheless, the invention is characterised in that no sample is allocated to a certain class, but it is classified in a data record based on the interpolation between different classes.

(57) Zusammenfassung: Die Erfindung bezieht sich auf das Gebiet von Analysemethoden, die zur Diagnose von Krankheiten an menschlichen und tierischen Proben geeignet sind. Die Erfindung offenbart ein Auswerteverfahren, durch das es möglich wird, einzelne Krankheitsstadien zu diagnostizieren, so dass der Verlauf einer Krankheit abgebildet werden kann. Hierdurch können eine Früherkennung von Krankheiten sowie Therapiekontrollen etc. geleistet werden. Das erfindungsgemässe Verfahren macht sich multivariable Auswerteverfahren zunutzen, die im Stand der Technik zur Klassifizierung von Proben bekannt sind. Anders als im Stand der Technik erfolgt jedoch erfindungsgemäss keine Zuordnung einer Probe zu einer Klasse, sondern eine Einordnung in einen Datensatz, der sich aus der Interpolation zwischen verschiedenen Klassen ergibt.

WO 2004/090779 A2



(81) **Bestimmungsstaaten** (soweit nicht anders angegeben, für jede verfügbare nationale Schutzrechtsart): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) **Bestimmungsstaaten** (soweit nicht anders angegeben, für jede verfügbare regionale Schutzrechtsart): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Erklärungen gemäß Regel 4.17:

- hinsichtlich der Berechtigung des Anmelders, die Priorität einer früheren Anmeldung zu beanspruchen (Regel 4.17 Ziffer iii) für die folgenden Bestimmungsstaaten AE, AG,

AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW, ARIPO Patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

— Erfindererklärung (Regel 4.17 Ziffer iv) nur für US

Veröffentlicht:

- ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

KRANKHEITSVERLAUFKONTROLLE

Die Erfindung bezieht sich auf das Gebiet von Analysemethoden, die zur Diagnose von Krankheiten an menschlichen oder tierischen Proben geeignet sind. Zur Analyse der Proben werden elektromagnetische Spektren aufgenommen und gespeichert und mittels multivariater Auswerteverfahren analysiert. Die Diagnose von Krankheiten mittels multivariater Auswerteverfahren ist im Stand der Technik bereits bekannt und wird im Nachfolgenden noch näher erläutert. Mit Hilfe des erfindungsgemäßen Verfahren ist es jedoch erstmals möglich, neben der Erkennung einer spezifischen Krankheit auch einzelne Verlaufsformen eines Krankheitsbildes sicher zu erkennen. Das erfindungsgemäße Verfahren ermöglicht somit erstmals einzelne Verlaufsstadien, eine Krankheit zu diagnostizieren, die zum Beispiel Hinweis auf die Entstehung einer Krankheit geben können, oder Auswirkung von Therapiemethoden auf den Gesundheitszustand eines Patienten widerspiegeln.

Im Stand der Technik beruhen die meisten diagnostischen Verfahren derzeit auf der Erkennung einzelner charakteristischer Stoffe in einer zu untersuchenden Probe, die Hinweis auf eine Krankheit geben. Diese Stoffe werden zum Beispiel im Blut, Urin oder andere Körperflüssigkeiten nachgewiesen. Für eine Krankheit charakteristische Stoffe sind z. B. krankheitsbedingte Stoffwechselprodukte, die sich im Körper akkumulieren und im Blut nachgewiesen werden können. Die von den Patienten entnommenen Proben werden mittels klinischer Analyseverfahren oder spektroskopischer Verfahren untersucht und in einem Zusammenhang mit den Krankheitssymptomen des Patienten gestellt. Aufgrund dieser Befunde erfolgt eine ärztliche Diagnose, die sowohl die Analyseergebnisse der Probe als auch die Krankheitssymptome des Patienten berücksichtigt.

Oftmals zeigt es sich jedoch, dass die Diagnose einzelner Krankheiten besonders schwierig ist, da hier die für die Krankheit zu analysierenden Stoffe nicht eindeutig bestimmbar oder sogar unbekannt sind. Eine Diagnose derartiger Krankheiten kann häufig nur durch aufwendige Analysen vielfältiger Parameter gewährleistet werden, die einen Hinweis auf die Krankheit geben. Ein Beispiel für eine schwer diagnostizierbare Krankheit

ist z. B. das metabolische Syndrom, unter dem – wie vermutet wird - in den Vereinigten Staaten mindestens jeder 10. Erwachsene leidet. Damit ist diese Erkrankung eine der häufigsten Gesundheitsanomalien in einer westlichen Zivilisation. Beim metabolischen Syndrom handelt es sich um eine Ansammlung von fünf verschiedenen Risikofaktoren, die eine hohe kardiovaskuläre Sterblichkeit nach sich ziehen. Ein metabolisches Syndrom gilt dann als diagnostiziert, wenn mindestens drei von fünf Risikofaktoren vorhanden sind. Zu den Risikofaktoren bzw. Parametern, die zur Bestimmung dieser Krankheit zählen, gehören Fettsucht, Bluthochdruck, Cholesterinwerte sowie erhöhter Blutzucker- und Blutfettwert. Das angeführte Beispiel zeigt, dass zur Diagnose dieser Krankheit fünf unterschiedliche Untersuchungen notwendig sind, die dann auf das Krankheitsbild schließen lassen.

Um derartige Aufwendige Untersuchungsverfahren zu vereinfachen und eine größere Diagnosesicherheit zu gewinnen, wurden Verfahren zur Krankheitsanalyse entwickelt, die auch ohne die Kenntnis von einem eindeutigen für eine Krankheit charakteristischen Stoff eine sichere Diagnose mit nur einem Messverfahren zulassen. Diese Verfahren werden z. B. mittels der IR-Spektroskopie durchgeführt und machen sich u. a. multivariate Auswerteverfahren zunutzen. Mit Hilfe multivariater Auswertealgorithmen ist es dabei möglich, eine Mustererkennung von Spektren vorzunehmen, ohne dass die Charakterisierung einzelner Stoffe vorgenommen werden muss. Multivariate Auswerteverfahren erlauben dabei, mit einer hohen Zuverlässigkeit eine unbekannte Probe zu analysieren und sie in Abhängigkeit von einer zuvor trainierten Datenbank verschiedenen Krankheiten zuzuordnen. Zur Erstellung der Datenbank werden zunächst Proben von Patienten entnommen, die ein spezifisches Krankheitsbild aufweisen. Die vermessenen Spektren werden anschließend einer Klasse zugeordnet, die die jeweilige Krankheitsfamilie charakterisiert. Wird für eine Vielzahl von Proben Infrarotspektren nach identischen oder zumindest vergleichbaren Messverfahren aufgenommen, die einer bekannten Klasse angehören, kann ein Training der Datenbank erfolgen. Diese Spektren werden entsprechend ihrer Klassifikation dem multivariaten Auswerteverfahren zugeführt, sodass die Datenbank charakteristische Spektrenmuster, die jeweils einer Klasse angehören und somit einer Krankheitsfamilie zugeordnet sind, beinhaltet. Mit Hilfe von multivariaten Auswertealgorithmen wird das System solange trainiert, bis zwischen verschiedenen Klassen unterschieden werden kann. Beispielsweise erfolgt so eine Kategorisierung von Proben, die von gesunden bzw. an einer zu untersuchenden Anomalie erkrankten Menschen entnommen wurden. Wird nun eine unbekannte Probe ver-

messen, kann das System mit hinreichender Zuverlässigkeit die unbekannte Probe einer Klasse zuordnen, sodass eine Diagnosehinweis gegeben wird. Hierbei findet vorzugsweise eine Mustererkennung des vermessenen Spektrums in einem zuvor definierten, vermessenen Wellenlängenbereich statt, auf dessen Basis eine Einordnung einer unbekannten Probe in eine zuvor klassifizierte Krankheitsgruppe (krank/gesund) erfolgt.

5 Grundprinzipien eines derartigen Verfahrens, das eine Klassifizierung von Proben mittels IR-Spektroskopie und multivariaten Auswerteverfahren ermöglicht, sind z. B. in dem Dokument EP 0 644 412 beschrieben.

Des weiteren wird in dem Dokument EP 02 000 952.8 die Anwendung eines derartigen

10 Verfahrens zur Analyse des metabolischen Syndroms offenbart. Hierfür werden in einer Datenbank, die in einem vorbestimmten Wellenlängenbereich Spektren von gesunden Patienten sowie Spektren von Patienten, die unter dem metabolischen Syndrom leiden, gespeichert und klassifiziert.

In Abhängigkeit der zu erstellenden Analyse wird ein Wellenlängenbereich ausgewählt,

15 in dem charakteristische Spektrenmuster eine Zuordnung zu einer Klasse sicher erlauben. Eine Bestimmung des jeweiligen Wellenlängenbereiches, der sich zur Analyse eines bestimmten Krankheitsbildes als geeignet erweist, kann z. B. mittels multivariater Verfahren erfolgen oder Algorithmen, wie sie z. B. in „NMA Biomedicine“, Vol. 11, Issue 4 - 5 (1998), pp. 209-216, beschrieben sind. Bei einer Vielzahl von Auswerteverfahren, wie

20 der Diskriminanz- und Clusteranalyse, wird folglich zunächst eine Korrelationsmatrix aufgestellt. Aufgrund der Korrelationsmatrix ermittelt das Auswerteverfahren welche Wellenlängen mit welchem Einfluss zur Klassifizierung verwendet werden sollen. Die zur Auswertung herangezogenen Wellenlängen spannen einen Raum auf, der in seiner Dimension mit der Zahl der Wellenlängen übereinstimmt. Von dem Auswerteverfahren

25 wird ein Unterraum gewählt, in dem die Zuordnung von Proben zu den Klassen möglichst eindeutig erfolgt. Hierbei ergibt sich die Zuordnung von Proben zu einer Klasse aus ihrer räumlichen Nähe in dem genannten Unterraum. Sobald ein Schema zur Klassifikation entsprechend einer Krankheitsfamilie entwickelt wurde, können die Daten im System gespeichert werden, sodass ein Analysesystem Diagnosehinweise für eine

30 Vielzahl von Krankheiten leisten kann.

Im Folgenden werden anhand des Beispiels der Diskriminanzanalyse multivariate Auswerteverfahren, wie sie im Stand der Technik bekannt sind, näher erläutert.

Die lineare Diskriminanzanalyse ermöglicht eine Trennung zwischen zwei Datensätzen, die im aufgeführten Beispiel, z. B. Patientendaten eines gesunden und eines erkrankten Patienten, widerspiegeln, mit dem Ziel eine sichere Unterscheidung zwischen den Datensätzen zu ermöglichen. Hierbei zeigt die Praxis, dass es häufig zu Überlappungen der Datensätze kommt, wobei eine Anzahl von Daten sowohl den Datensatz „gesund“ als auch den Datensatz „krank“ zugeordnet werden kann. Eine eindeutige Klassifizierung dieser Daten, die im angeführten Beispiel als Schnittmenge der Datensätze „gesund“ und „krank“ bezeichnet werden kann, ist somit nicht gegeben. Eine Zuordnung von unbekannten Probe, deren vermessene Parameter in diesen Schnittbereich der beiden Datensätze liegen, ist deshalb nicht möglich. Die lineare Diskriminanzanalyse zieht analog zu anderen multivariaten Auswerteverfahren eine Grenze zwischen zwei Datensätzen, zu denen eine Zuordnung jeweils erfolgen soll. Alle Daten, die innerhalb dieser Grenze/Grenzwerte liegen, gehören z. B. der Klasse „gesund“ an. Je größer der Datensatz einer Klasse gewählt wird, desto sensibler ist eine Zuordnung einer unbekannten Probe in diese Klasse, sodass nahezu alle Proben, die dieser Klasse angehören, auch dieser zugeordnet werden. Ein derartig großer Datensatz umfasst dann jedoch auch die Zuordnung der Proben, die der Klasse fälschlicher Weise zugeordnet wurden, sodass eine sensitive Zuordnung eine entsprechend geringe Spezifität aufweist. Aufgabe von derartigen Auswerteverfahren ist es deshalb, ein Optimum der Zuordnung zu erreichen, das zwischen Sensitivität und Spezifität abwägt. Dabei ist ein Training des Programms, das Sensitivität und Spezifität der Zuordnung erhöht mittels der Vermessung und Zuordnung von bekannten Proben möglich. In der Praxis werden bei der Auswertung die Verfahrensparameter entsprechend der jeweiligen Zuordnung angegeben, die die Sensitivität und Spezifität des Verfahrens kennzeichnen. Hierbei werden die Werte von unbekannten Proben, die sich unmittelbar in einer Schnittmenge von Datensätzen befinden, als nicht auswertbar gewertet, um einen zu großen Fehler bei der Zuordnung zu vermeiden. Ein Diagnosehinweis zu einer Probe, deren Parameter innerhalb der Schnittmenge liegen, ist nicht möglich.

Nachteil des Standes der Technik ist folglich, dass die Diagnose einer Krankheit nur bezüglich verschiedener Klassen (z. B. „gesund“ oder „krank“) erfolgen kann, ohne dass z. B. Daten innerhalb der Schnittmenge berücksichtigt werden können. Ein an sich kontinuierlicher Krankheitsverlauf eines Patienten wird somit in zwei diskrete Schritte eingeteilt, die als Diagnosehinweis ausschließlich den Zustand „gesund“ oder „krank“

zulassen. Ein realer Verlauf einer Krankheit mit verschiedener Krankheitsstadien kann somit nicht berücksichtigt werden.

Das erfindungsgemäße Verfahren hat die Aufgabe, die beschriebenen Nachteile des Standes der Technik zu vermeiden. Insbesondere berücksichtigt dabei das erfindungsgemäße Verfahren den kontinuierlichen Verlauf einer Krankheit und ermöglicht somit insbesondere auch eine Auswertung von Spektren innerhalb einer Schnittmenge von Datensätzen. Das erfindungsgemäße Verfahren macht sich folglich gerade diese Daten zunutze, die nach dem Stand der Technik als nicht auswertbar gelten. Durch das erfindungsgemäße Verfahren wird z. B. eine Früherkennung von Krankheiten sowie eine Diagnostizierung von Krankheitsverläufen etc. ermöglicht.

Die Erfindung beinhaltet ein Verfahren zur Auswertung eines elektromagnetischen Spektrums, wobei eine Zugehörigkeit des elektromagnetischen Spektrums zu einer Klasse anteilmäßig bestimmt wird. Auf diese Weise wird gewährleistet, dass ein realer Gesundheits-/bzw. Krankheitsverlauf abgebildet werden kann, sodass wesentliche Rückschlüsse z. B. auf eine Therapie getroffen werden können oder krankheitsvorbeugende Maßnahmen möglich werden.

Das erfindungsgemäße Verfahren analysiert ein elektromagnetisches Spektrums, das durch eine Vermessung einer unbekannten Probe generiert wurde. Mittels eines multivariaten Auswerteverfahrens wird hierbei ein Probenparameter ermittelt, der zu einem Datensatz zugeordnet werden kann. Der Datensatz wird durch die Interpolation zwischen mindestens einem Teil von ersten Daten und mindestens einem Teil von zweiten Daten berechnet, wobei die ersten und zweiten Daten jeweils einer ersten und zweiten Klasse zugeordnet sind. Hierbei beinhalten die ersten Daten einen ersten Parameter, der anhand von elektromagnetischen Spektren von bekannten Proben, die der ersten Klasse angehören, mittels des multivariaten Auswerteverfahrens generiert wurde, sodass eine zuverlässige Zuordnung von Proben zu der ersten Klasse aufgrund des ersten Parameters möglich ist.

Die zweiten Daten beinhalten analog einen zweiten Parameter, der ebenfalls anhand von elektromagnetischen Spektren von bekannten Proben, die der zweiten Klasse angehören, mittels des multivariaten Auswerteverfahrens generiert wurde, sodass eine zuverlässige Zuordnung von Proben zu der zweiten Klasse aufgrund des zweiten Parameters möglich ist.

Nach der Zuordnung des Probenparameters zum Datensatz, wird ein Anteilswert des Probenparameters an zumindest einem Teil des Datensatzes bestimmt. Aufgrund der Basis dieses Anteilswertes wird ein Anteil des Probenparameters an einer Klasse wiedergegeben, sodass der Anteil der unbekannten Probe bestimmt wird, mit dem die
5 unbekannte Probe einer Klasse angehört.

Das erfindungsgemäße Verfahren ermöglicht eine kontinuierliche Klassifizierung einer Probe zwischen verschiedenen Krankheitsfamilie. Aufgrund der Interpolation zwischen Daten, die z. B. im Spätstadium einer Krankheitsfamilie gefunden werden und Daten, die von einem gesunden Patienten ermittelt werden können, kann der Verlauf einer
10 Krankheit durch eine Verlaufskurve der interpolierten Daten abgebildet werden. Im Rahmen der Erfindung wird der interpolierte Datensatz durch einen Datenbereich definiert, der durch die Interpolation zwischen mindestens zwei Werten aufgespannt wird, so dass der Datensatz mindestens einen Zwischenwert beinhaltet, der den aufgespannten Datenbereich zwischen den bekannten Werten definiert. Hierbei gehören
15 die bekannten Daten, zwischen denen eine Interpolation erfolgt, erfindungsgemäß jeweils der ersten und zweiten Klasse an.

Eine Zuordnung von vermessenen Daten zu einer Verlaufskurve des interpolierten Datensatzes ermöglicht, das Krankheitsstadium einer Probe genau abzubilden und so z. B. eine Prognose über den Verlauf der Krankheit zu erstellen oder ggf. eine
20 Wahrscheinlichkeit einer möglichen Erkrankung zu berechnen. Eine derartige Einordnung der Probe bietet somit u. a. die Möglichkeit, Frühdiagnose sicher zu gewährleisten, die sich zur Heilung vieler Krankheiten als wesentlich erweisen. Die kontinuierliche Differenzierung des Krankheitsbildes erlaubt hierbei eine Verlaufsprognose für Krankheiten und ein Monitoring der jeweiligen akuten
25 Verlaufsform abzubilden, ohne dass das Verfahren auf die ausschließliche Zuordnung zwischen zwei Klassen beschränkt ist. Eine Auswertung von Daten, die nach dem Stand der Technik innerhalb einer Schnittmenge liegen, wird möglich.

Das erfindungsgemäße Verfahren befasst sich mit der Analyse von Körperflüssigkeiten, Zell- und Bakteriensuspensionen. Klinisch relevante Körperflüssigkeiten, die mit diesem
30 Verfahren ausgewertet werden können, sind z. B. Blut, Plasma, Serum, Hämolysat, Urin, Liquor, Speichel, Transpirationsflüssigkeit, Augenflüssigkeit, Zystenflüssigkeit, Fruchtwasser, Sputum, Ascitis, Sperma. Unter Zell- und Bakteriensuspensionen sind solche

Proben zu verstehen, die entweder direkt (z. B. durch Punktion, Abstrich etc.) oder durch Anzucht in einem Nährmedium gewonnen werden.

Insbesondere ist ein erfindungsgemäßes Verfahren geeignet, Flüssigkeiten mit einer Vielzahl verschiedener Inhaltsstoffe zu analysieren. Dies gilt besonders für Blut, Plasma,
5 Serum und Hämolysat.

Die Durchführung des erfindungsgemäßen Verfahrens beginnt mit einem Zurverfügungstellen der zu klassifizierenden Probenflüssigkeit bzw. -suspension. Dies kann in ausreichend bekannter Weise z. B. durch die Entnahme einer Blut- oder Urinprobe erfolgen. Für einige Arten von Proben ist es vorteilhaft, wenn vor der Durchführung des
10 Verfahrens eine Aufarbeitung erfolgt. Vollblut kann beispielsweise zentrifugiert werden, um Serum und Blutkuchen voneinander zu trennen.

Die gewonnene Probe kann vorteilhafterweise auf einen Träger aufgegeben werden, der es ermöglicht, ein Infrarotspektrum aufzunehmen. Ein solcher Träger kann beispielsweise ein metallisierter Objektträger sein. Bevorzugt sind solche Träger, die auf ihrer
15 Oberfläche reflektierend sind und eine Rauigkeit besitzen, die eine diffuse Streuung von Infrarotstrahlung bewirkt. Ebenfalls können solche Träger verwendet werden, die auf ihrer Oberfläche, auf die die Probe aufgebracht wird, verspiegelt sind. Eine Beschichtung der Träger kann mit Metallen, vorzugsweise Edelmetallen wie Gold, Platin, Palladium aber auch z. B. mit unedlen Metallen, vorzugsweise Aluminium und
20 Chrom erfolgen. Diese Beschichtungen besitzen den Vorteil, dass sie weitgehend inert gegenüber der Probe und auch gegenüber der Umgebungsluft sind. Die Träger können eine ebene Oberfläche besitzen, auf welche die Probe aufgegeben wird. Erfindungsgemäß bevorzugt sind jedoch solche Träger, die eine Vertiefung zur Aufnahme der Probenflüssigkeit besitzen.

25 Geeignete Probenträger werden z. B. in dem Patent US 5,734,587 beschrieben, für Transmissionsmessungen sowie für Reflektionsmessungen in dem Patent US 5,869,001.

Weitere bevorzugte Ausführungsformen von Probenträgern werden z. B. in dem Dokument EP 0 644 412 und DE 4331596 beschrieben.

Neben zur Reflektionsmessung geeigneten Trägern können auch Träger zur Transmissionsmessung vorteilhaft verwendet werden, welche die eingestrahlte Strahlung hin-
30 durchlassen. Die genannten Träger können zum Beispiel aus Kunststoffen, wie Poly-

carbonat, Polyethylen oder Polymethylmethacrylat, bestehen, die für Infrarotstrahlung in weiten Wellenzahlbereichen durchlässig sind. Diese Träger für die Transmissionsmessung können analog denen für die Reflektionsmessung Vertiefungen für die Probenflüssigkeit besitzen. Hierbei besitzen Messungen in diffuser Reflektion den Vorteil, dass Inhomogenitäten des Probenmaterials, z. B. Trübungen, die Auswertung nicht stören, sondern sogar die diffuse Reflektion unterstützen. Zur Verwendung der oben genannten Träger hat es sich als vorteilhaft erwiesen, die aufgetragene Probe vor dem Vermessen zunächst eintrocknen zu lassen. Beispiele zum Trocknen der Probe werden z. B. in dem US-Patent 5,869,001 offenbart. Es besteht jedoch auch die Möglichkeit, auf den Trockenschritt zu verzichten und die flüssige Probe direkt zu vermessen. Messungen an flüssigen Proben werden beispielsweise im Dokument WO 02/057753 beschrieben. Hierbei erfolgt die Messung in Transmission durch eine Messzelle mit einer Dicke von wenigen μm .

Des weiteren erweisen sich auch Messungen unter abgeschwächter Totalreflektion (ATR-Messungen) für das erfindungsgemäße Verfahren als geeignet.

Anordnungen zur Auswertung diffuser oder direkter Reflektion sowie zur Transmissionsmessung sind im Stand der Technik prinzipiell bekannt und werden deshalb an dieser Stelle nicht näher erläutert. Eine Anordnung, die zur Auswertung diffus reflektierter Strahlung dient, ist z. B. in der Patentanmeldung DE 43 31596 detaillierter dargestellt.

Prinzipiell gibt es zwei verschiedene Messtechniken, die häufig in der Spektroskopie zur Untersuchung von Proben im diagnostischen Bereich verwendet wird. Hierbei wird bei der Infrarotspektroskopie häufig im mittleren Infrarotbereich ($2,5 - 25 \mu\text{m}$) gemessen. Zur Aufnahme eines Infrarotspektrums wird die Probe auf dem Probenträger mit Infrarotstrahlung beleuchtet. Infrarotstrahlung in den bereits beschriebenen Wellenzahlbereichen kann mit im Markt befindlichen IR-Spektrometern von beispielsweise Nicolet, Bruker und Perkin Elmer erzeugt werden.

In Abhängigkeit von der zu untersuchenden Probe verwendet man ebenfalls häufig die Raman-Spektroskopie, die jedoch im Vergleich zur Infrarotspektroskopie eine wesentlich geringere Sensitivität zum Nachweis eines Analyten aufweist. Hierfür macht man sich jedoch gerade bei biologischen Proben häufig den Vorteil zunutze, dass, anders als in der Infrarotspektroskopie, nur geringfügig störende Wasserbanden das Analytsignal

überlagern. Gerade bei biologischen Proben, bei denen der Wassergehalt häufig sehr hoch ist, führen Infrarotmessungen häufig zu verfälschten Analyseergebnissen, da eine starke Absorptionsbande von Wasser im mittleren Infrarotbereich liegt. Trotz der schlechteren Sensitivität erweist sich somit die Raman-Spektroskopie unter Umständen als eine geeignete Spektroskopiemethode. Nähere Informationen zu den genannten Spektroskopiemethoden sowie möglichen Anwendungen werden z. B. in dem Dokument „Mid-IR and Raman spectroscopy for medical diagnostics“, Appl. Spectros. Rev. 36, pp. 181 – 237 (2001), bei W. Petrich, vorgestellt.

Um die Absorption von Wasser bei der Infrarotspektroskopie weitestgehend zu minimieren, kann es sich als vorteilhaft erweisen, die Probe vor dem Vermessen, wie bereits beschrieben, zunächst auf dem Probenträger eintrocknen zu lassen. Dabei hat sich herausgestellt, dass das Eintrocknen in der Regel zu keiner Beeinträchtigung der Probe führt, die eine Analyse und Klassifikation unmöglich macht. Die Eintrocknung ist hingegen ein wichtiger Schritt, um die benötigte hohe Reproduzierbarkeit bei der Spektrenaufzeichnung zu erhalten.

Das Eintrocknen der Probe kann durch einfaches Verdampfen der Flüssigkeit bei Raumtemperatur oder aber auch unter besonderen Bedingungen, beispielsweise durch Anlegen eines Vakuums, Erwärmung der Probe oder Aufbewahrung in Gegenwart eines Trockenmittels, erfolgen. Hierbei wird unter dem Begriff Eintrocknen im Wesentlichen keine hundertprozentige Trocknung der Probe gemeint, sodass eine getrocknete Probe wenige Prozent Restfeuchte enthält. Die Trockenschritte – wie beschrieben – besitzen dabei den Vorteil, dass die Probe einen zusammenhängenden Film ohne Risse ausbildet, die eine Auswertung beeinträchtigen könnten.

Beim Eintrocknen der Probe entsteht abhängig von den Trockenbedingungen, Träger und der Art der Probe ein Trocknungsrest von variierender Gestalt. Beim Eintrocknen von Serumproben auf geschlossenen Oberflächen bildet sich in der Regel ein kreisförmiger Wulst, dessen Dicke in Richtung auf das Kreiszentrum hin abnimmt. Bei freitragenden Filmen bildet sich in den Randzonen zum Trägermaterial ebenfalls ein Wulst, der zur Mitte der Öffnung in einen relativ gleichmäßig dicken Film übergeht.

Die Bestrahlung der Probe in Reflexion und Transmission erfolgt bevorzugt so, dass die gesamte Probe von dem IR-Strahlenbündel erfasst wird. Aufgrund dieser integralen Bestrahlung der Probe ist die Messung weitestgehend unabhängig von dem Trock-

nungsprofil der Probe. Dies ist ein weiterer, wesentlicher Faktor, der die Reproduzierbarkeit der Spektrenauswertung erhöht.

Wird Raman-Spektroskopie als Analysemethode verwendet, ist ein Trockenschritt während der Probenpräparation nicht notwendig.

- 5 Zur Aufnahme eines Infrarotspektrums wird die Probe mit Infrarotstrahlung beleuchtet. Die Infrarotstrahlung kann innerhalb des gesamten Infrarotbereichs, d. h. etwa 400 bis 15.000 cm^{-1} liegen. Besonders vorteilhaft in Zusammenhang mit der vorgenannten Trocknungsmethode ist die IR-Spektroskopie im Bereich von 400 bis 4.000 cm^{-1} , da sie in diesem Bereich eine hohe Spezifität besitzt.
- 10 Die durch die Probe hindurchgetretene oder von der Probe reflektierte Strahlung wird durch einen Detektor aufgefangen, der einen möglichst großen Teil der Strahlung erfassen sollte. Im Stand der Technik sind infrarotsensitive Detektoren, wie z. B. Bolometer, Thermoelemente, photovoltaische Elemente und Halbleiterdetektoren bekannt. Bevor die Strahlung in den Detektor gelangt, kann sie durch ein System aus
- 15 Linsen, Blenden und Spiegeln geleitet werden.

Die Aufnahme von Infrarotspektren ist aus dem Stand der Technik hinlänglich bekannt. Bezüglich der Aufnahme von IR-Spektren eingetrockneter Proben sei hier auf die Patentanmeldung DE 4331596 verwiesen.

- Bei den üblichen Verfahren der Spektrenauswertung wird in Abhängigkeit von der eingestrahnten Wellenlänge die transmittierte oder reflektierte Strahlung aufgezeichnet.
- 20 Erfindungsgemäß wird ein solches Spektrum durch ein multivariates Auswerteverfahren zur Genierung eines Probenparameters ausgewertet. Im Rahmen der Erfindung umfasst dabei der Begriff Spektrum jegliche Form der Datenstruktur, die durch die Vermessung einer Probe, wie beschrieben, generiert wird. Dabei ist es für das erfindungsgemäßen
- 25 Auswerteverfahren ohne Bedeutung, ob dem Verfahren z. B. die Daten eines Interferogramms oder die hieraus resultierenden fouriertransformierten Daten zugrunde gelegt werden. Für diesen Auswerteprozess muss das Spektrum zunächst digitalisiert werden, sofern es nicht bereits digital aufgezeichnet wurde. Dies bedeutet, dass das Spektrum in eine Wertetabelle umgesetzt wird, die für diskrete Wellenzahlen die
- 30 Absorption oder Transmission enthält. Die Zahl der Wertepaare richtet sich im wesentlichen nach der Kapazität der vorhandenen Soft- bzw. Hardware. Erfindungsgemäß ist es nicht notwendig, aus dem Spektrum einzelne Banden zu analysieren und getrennt

auszuwerten. Es hat sich für einige Analysen jedoch als vorteilhaft erwiesen, wenn die Wertetabelle, die zur Auswertung kommt, auf bestimmte Bereiche der Wellenzahlen beschränkt wird. Prinzipiell können vielfältige Verfahren zur Selektion von Wellenlängenbereichen verwendet werden. Derartige Verfahren sind im Stand der Technik
5 hinlänglich bekannt und z. B. in EP 0 644 412 und WO 02/66963 dargelegt. Auf diese Weise kann eine Auswahl relevanter Bereiche der Spektren erfolgen und Spektrenteile, die von geringem Informationsgehalt sind, können eliminiert werden.

Das digitalisierte Spektrum wird einem multivariaten Auswerteverfahren, das als Computerprogramm zur Verfügung steht, übergeben.

- 10 Multivariate Auswerteverfahren sind im Sinne der Erfindung Diskriminanzanalysen, regularisierte Diskriminanzanalysen, neuronale Netze, Clusteranalysen usw. Prinzipiell sind jegliche denkbaren multivariaten Auswerteverfahren, wie sie im Stand der Technik hinreichend bekannt sind, bei der Durchführung des erfindungsgemäßen Verfahrens möglich.
- 15 Genannte Softwarepakete können beispielsweise von den Firmen STSC, STAT SOFT, SAS und Galactic Industries bezogen werden.

- Die multivariaten Auswerteverfahren müssen vor ihrer erfindungsgemäßen Anwendung auf das Anwendungsgebiet trainiert werden. Dies erfolgt in der Regel durch eine Zahl von Trainingsläufen, bei denen Infrarotspektren einer Mehrzahl von Proben bekannter
20 Klassifikationen in das Auswerteverfahren eingegeben werden. Ebenfalls wird die Art der Klassifikation der einzelnen Proben dem Auswerteverfahren mitgeteilt. Bei dem genannten Trainingslauf wählt der Auswertalgorithmus selbstständig Parameter, die eine sichere Zuordnung der Proben zu den vorgegebenen Klassen sicherstellen. Die auf diese Weise gewonnenen Parameter können im späteren Verlauf zur Berechnung des interpolierten Datensatzes dienen. Hierbei ist es erfindungsgemäß ohne Bedeutung, ob eine
25 Interpolation zwischen den Parametern direkt erfolgt oder eine Interpolation zunächst zwischen den Spektrendaten durchgeführt wird, bevor eine Generierung der Parameter erfolgt. Es zeigt sich jedoch, dass zunächst die Bestimmung der Parameter und anschließend eine Interpolation zwischen den Parametern eine Generierung des interpolierten
30 Datensatzes erheblich vereinfacht.

Das Analysesystem kann in der Regel eine Vielzahl von verschiedenen Klassen beinhalten, zwischen denen erfindungsgemäß interpoliert werden kann. Das erfindungsgemäße

Verfahren ermöglicht durch die Interpolation zwischen Daten einen Krankheitsverlauf widerspiegeln zu können und somit auch Zwischenstufen eines Krankheitsbildes genau abzubilden und zu verfolgen. Mit Hilfe des erfindungsgemäßen Verfahrens erfolgt eine Einordnung der Probe in verschiedene Verlaufskurven unterschiedliche Krankheitsverläufe, die mit bisherigen Methoden eine Vielzahl separater Analyseprozesse benötigte. Hierbei sind Klassen, die in den Bereich der inneren Krankheiten fallen, für die klinische Analyse besonders wichtig. Beispielsweise ist mit erfindungsgemäßen Verfahren die Generierung von Verlaufskurven möglich, die es erlaubt, vorausszusagen, ob eine Leber- oder Nierenkrankheit vorliegt. Die beschriebene Einordnung einer unbekannten Probe ist jedoch nicht mit einer eindeutigen Diagnose gleichzusetzen, da gleiche spektrale Veränderungen der Probe unter Umständen durch verschiedene Krankheiten oder Störungen hervorgerufen werden können. Die genannte Einordnung der Probe stellt vielmehr ein Screeningverfahren dar, mit dem entschieden werden kann, auf welche Krankheiten die Probe mit klassischen Verfahren geprüft werden soll. Hieraus resultiert eine Einsparung einer Vielzahl von analytischen Tests, die normalerweise, z. B. bei einer Eingangskontrolle in Kliniklabors, notwendig wären.

Darüber hinaus ermöglicht das erfindungsgemäße Verfahren die Früherkennung von Krankheiten, die unter Umständen mittels klassischer Verfahren nicht geleistet werden können.

Wie im weiteren Verlauf noch dargestellt wird, ist es mittels des erfindungsgemäßen Verfahrens gelungen, erstmals Zusammenhänge zwischen verschiedenen Krankheiten abzubilden, die bislang unabhängig voneinander betrachtet und behandelt wurden. Es wird schnell ersichtlich, dass derartige neu gewonnene Erkenntnisse über Zusammenhänge von bislang als unabhängig betrachteten Krankheiten, vielfältige neue Möglichkeiten der Therapie und Vorsorge bieten.

Erfindungsgemäß kann die Interpolation zwischen zwei Klassen (z. B. gesund und krank) am Beispiel der linearen Diskriminanzanalyse erfolgen, indem vorteilhafterweise der Abstand zwischen der Diskriminanzfunktion und des Probenparameters der zu untersuchenden Probe im multidimensionalen Raum berechnet wird. Dieser Raum könnte z. B. der Raum der Hauptkomponenten sein, der Abstand könnte definiert werden durch den euklidischen Abstand, oder dem Mahalanobis-Abstand. Dieses Abstandsmaß kann dann über den gesamten Datensatz des Trainings normiert werden und als besondere Ausformung einer Interpolation verstanden werden. Alternativ hierzu könnte

der Abstand mit Hilfe einer Wahrscheinlichkeitsfunktion umgerechnet werden. Ebenso ist es denkbar, den Wert der Diskriminanzfunktion zu benutzen, oder einen mit Hilfe einer Wahrscheinlichkeitsfunktion modifizierten Wert der Diskriminanzfunktion.

Im Fall einer Auswertung durch neuronale Netze ist es denkbar, die Gewichtung der
5 jeweiligen Vernetzung bzw. Neuronen zu den beiden Ausgangsneuronen, z. B. den Klassen krank und gesund, miteinander zu kombinieren, zu normieren und eine Skalierung dieser normierten Gewichtung zwischen den beiden Klassen/ Ausgangsneuronen als Interpolation zu benutzen.

Auch hier ist die Umformung mit einer beliebigen Funktion denkbar, z.B. einer Wahr-
10 scheinlichkeitsfunktion. Bei einer überwachten Klassifikation wie sie bei einem neuronalen Netz der Fall ist, ist es ferner möglich beim Trainieren des Netzes eine (oder mehrere) neue Klasse – vertreten durch ein oder mehrere neues Ausgangsneuron – zu implementieren und hier solche Spektren zu verwenden, die nicht hinreichend der Gruppe A oder B (z.B. gesund oder krank) zugeordnet werden können.

15 Im Folgenden wird eine Interpolation zwischen Daten anhand des Krankheitsbildes, Diabetes, beispielhaft veranschaulicht. Eine Interpolation erfolgt zwischen ersten Daten, die einer Klasse angehören, die als „gesund“ bezeichnet wird, d. h. diese Daten wurden anhand von bekannten Proben, die von gesunden Patienten oder Patienten, die nicht an Diabetes erkrankt sind, entnommen wurden, generiert- und zweiten Daten, die anhand
20 von an Diabetes erkrankten Patienten entnommenen Proben generiert wurden. Basierend auf den jeweiligen Daten der Klasse werden zunächst mittels eines ausgewählten multivariaten Auswerteverfahrens Parameter generiert, sodass die Parameter jeweils der ersten und der zweiten Klasse zugeordnet werden können. Anschließend erfolgt eine Interpolation zwischen den Parametern. Hierbei zeigt sich vorteilhafterweise, dass ein
25 Verlauf einer Krankheit mit dem auf diese Weise berechneten Kurvenverlauf der interpolierten Daten korreliert, sodass Proben von Patienten in unterschiedlichen Krankheitsstadien dem interpoliertem Datensatz zugeordnet werden können. Wird ein Auswerteverfahren wie beschrieben durchgeführt, wird bei der Analyse einer unbekannten Probe zunächst mittels des selben multivariaten Auswerteverfahrens ein Probenparameter ermittelt, sodass eine Zuordnung des Probenparameters zu dem Datensatz, der
30 sich aus der Interpolation zwischen dem ersten und zweiten Parameter ergibt, erfolgen kann. Anhand dieser Zuordnung lässt sich nun ein Anteilswert des Probenparameters bestimmen, mit dem die Probe einer Klasse angehört. Beispielsweise kann der Anteils-

wert des Probenparameters an einer Klasse prozentual berechnet werden. Dieser Anteilswert spiegelt die Zuordnung des Probenparameters zu einem Punkt in der Verlaufskurve des interpolierten Datensatzes wider. Es ist weiterhin auch denkbar, dass der Anteilswert anstelle eines numerischen Wertes z. B. bildlich dargestellt wird. Eine

5 Zuordnung des Probenparameters zu der berechneten Verlaufskurve entsprechend einem Krankheitsverlauf veranschaulicht somit bildlich den Krankheitszustand des Patienten. Der Arzt erhält hierdurch einen Hinweis, in welchem Stadium der Krankheit sich der Patient, dessen Probe entnommen wurde, befindet.

Hierbei haben Untersuchungen ergeben, dass bei der Vermessung einer Probe eines

10 Patienten, der unter dem metabolischen Syndrom leidet, die hieraus generierten Parameter der Verlaufskurve der Krankheit Diabetes zugeordnet werden können. Eine Zuordnung des Probenparameters eines Patienten mit einem metabolischen Syndrom in die Verlaufskurve eines gesunden und an Diabetes erkrankten Patienten lässt somit darauf schließen, dass es sich bei dem metabolischen Syndrom um eine Vorstufe von Dia-

15 betes handelt. Es ergibt sich somit ein direkter Zusammenhang zwischen der Krankheit „metabolisches Syndrom“ und der Krankheit „Diabetes“. Das Beispiel verdeutlicht anschaulich, welche neuen Möglichkeiten durch das erfindungsgemäße Verfahren gegeben sind. Eine Vorbeugung der Krankheit Diabetes könnte somit bereits in einem frühen Stadium z. B. durch eine Medikamentenvergabe gegen Bluthochdruck, Fettsucht

20 oder erhöhten Cholesterinwerten etc. gewährleistet sein, sodass der Ausbruch der Krankheit Diabetes unter Umständen verhindert werden kann.

Die Erfindung beinhaltet weiterhin ein System, mit dem eine Zugehörigkeit einer Probe zu einer Klasse anteilmäßig bestimmt werden kann, sodass ein Gesundheitszustand charakterisiert wird. Das System beinhaltet eine Speichereinheit zur Speicherung der

25 von einem Detektor erhaltenen Signale eines elektromagnetischen Spektrums, das durch Vermessen einer unbekannten Probe generiert wurde. Das System beinhaltet weiterhin eine Auswerteeinheit zur Generierung eines Probenparameters mittels eines multivariaten Auswerteverfahrens auf der Basis der Signale des elektromagnetischen Spektrums. Mittels einer Auswerteeinheit wird der Probenparameter zu einem Datensatz

30 zugeordnet. Dieser Datensatz ergibt sich aus der Interpolation zwischen mindestens einem Teil von ersten Daten und mindestens einem Teil von zweiten Daten, die jeweils einer ersten und zweiten Klasse zugeordnet sind. Hierbei beinhalten die ersten bzw. zweiten Daten jeweils einen ersten bzw. zweiten Parameter, der jeweils der ersten bzw. der zweiten Klasse angehört. Diese Parameter wurden anhand von elektromagnetischen

- Spektren von bekannten Proben, die der jeweiligen Klasse angehören, mittels des multivariaten Auswerteverfahrens generiert, sodass eine zuverlässige Zuordnung von Proben zu der jeweiligen Klasse aufgrund des jeweiligen Parameters möglich ist. Vorteilhafterweise ist dieser Datensatz in einem permanenten Speicher der Speichereinheit
- 5 gespeichert. Es ist jedoch auch denkbar, dass die Speichereinheit einen Permanent-speicher enthält, in dem nur ein erster und ein zweiter Parameter, die jeweils einer ersten und zweiten Klasse angehören, gespeichert sind. Eine Berechnung des interpolierten Datensatzes erfolgt dann bei Bedarf. Die Auswertereinheit bestimmt weiterhin einen Anteilswert des Probenparameters an zumindest einem Teil des Datensatzes.
- 10 Dieser Anteilswert gibt einen Anteil des Probenparameters an einer Klasse wieder, sodass der Anteil der unbekannten Probe bestimmt wird, mit dem die unbekannte Probe einer Klasse angehört. Eine Ausgabereinheit zeigt die Zugehörigkeit einer Probe zu einer Klasse an. Vorteilhafterweise sind in dem System Daten zu mehr als zwei Klassen gespeichert, sodass ein System Diagnosehinweise für eine Vielzahl von unterschiedlichen
- 15 Krankheiten und Krankheitsverläufen leisten kann. In einer bevorzugten Ausführungsformen beinhaltet das System weiterhin eine Strahlenquelle zum Bestrahlen einer Probe mit elektromagnetischer Strahlung sowie einen Detektor, der die mit der Probe wechselwirkende Strahlung detektiert. Emittiert die Strahlenquelle z. B. elektromagnetische Strahlung im Infrarotbereich, wird vorteilhafterweise vom Detektor Signale
- 20 im Bereich von 1 μm bis 25 μm detektiert. Weiter vorteilhafte Ausführungsformen des Systems ergeben sich wie bereits beschrieben.

Im nachfolgenden werden anhand von Figuren Ausführungen der Erfindung näher erläutert, ohne dass hierdurch eine einschränkende Wirkung gegeben ist.

- 25 **Figur 1:** Spektrendaten einer Probe eines gesunden Menschen im Vergleich zu Proben eines an Diabetes erkrankten sowie an einem metabolischen Syndrom erkrankten Menschen.

Figur 2: Gegenüberstellung des DPR-Scores, der anhand von Proben eines gesunden, eines an metabolischem Syndrom erkrankten und eines an Diabetes erkrankten Menschen gewonnen wurde.

- 30 **Figur 3:** Parameterverteilung zweier Klassen

Figur 4: Balkendiagramm, das die Interpolation zwischen dem DPR-Score darstellt.

Figur 5: Braunwaldzyklus

Figur 1 zeigt verschiedene Auswertungen von Spektrendaten, die bei der Vermessung von Proben eines gesunden, eines am metabolischen Syndrom erkrankten Patienten und eines Diabetikers generiert wurden. Im Nachfolgenden werden die entsprechenden Proben und die diesbezüglichen Klassen der Einfachheit halber als gesund, metabolisches Syndrom und Diabetes bezeichnet.

In Figur 1a bis c werden jeweils charakteristische Wellenzahlenbereiche zwischen 900 und 1.200 cm^{-1} , 1.700 bis 1.800 cm^{-1} und 2.800 bis 3.000 cm^{-1} dargestellt. Hierbei stellt der mit 1 gekennzeichnete Kurvenverlauf in Figur 1 die Differenz zwischen den Daten gesund und Diabetes und die mit 2 gekennzeichnete Kurve zwischen gesund und metabolisches Syndrom dar. Kurve 3 veranschaulicht sich ergebene Differenzen der Spektren zwischen den Daten gesund und gesund, die durch zufälliger Unterschiede generiert werden.

Figur 1a zeigt den jeweiligen Kurvenverlauf in den genannten Wellenzahlenbereichen, die sich aus der Differenz der Mittelwerte der Datenpunkte jeweils ergeben. Um eine bessere Differenzierung zwischen den verschiedenen Zuständen erzielen zu können, werden im Stand der Technik bekannte statistische Auswertemethoden angewendet. Hierbei wird u. a. die Signifikanz von Unterscheidungskriterien in den Spektren beurteilt. Ein geläufiges Verfahren ist beispielsweise, die (quadrierten) Mittelwerte mit (der quadrierten Summe) den Standardabweichungen der jeweiligen Wellenlänge zu vergleichen.

Figur 1b zeigt die Anwendung des sogenannten Fischer-Kriteriums auf die entsprechend ausgewählten Wellenzahlenbereiche. Das Fischer-Kriterium erlaubt eine Gewichtung der sich ergebenden Differenzen in den Spektren. Hierzu werden die Abweichung zweier Kurven, die jeweils einer unterschiedlichen Klasse angehören, im Verhältnis zu der jeweiligen Standardabweichung innerhalb der Klasse darstellt. Es wird jedoch keine statistische Signifikanz der sich ergebenden Differenzen berücksichtigt.

Zur Berechnung der Signifikanz spektraler Unterschiede wird das im Stand der Technik als T-Test geläufige Auswerteverfahren angewendet, das die Anzahl der vermessenen Proben bei der Berechnung mit einbezieht. Figur 1c zeigt eine entsprechende Grafik, bei der der T-Wert gegen die genannten Wellenzahlenbereiche aufgetragen wird.

Wie dem Kurvenverlauf in den gezeigten Wellenzahlenbereichen zu entnehmen ist, ist mit Hilfe statistischer Auswerteverfahren eine Unterscheidung und somit eine eindeutige Klassifizierung von Proben in die Klassen gesund, metabolisches Syndrom und Diabetes möglich.

- 5 Figur 2 zeigt zu den in Figur 1 entsprechenden Daten der verschiedenen Klassen. Zur Klassifikation wurden die Daten mittels eines multivariaten Auswerteverfahrens ausgewertet, sodass jeweils ein Parameter generiert wurde, der der jeweiligen Klasse zugeordnet werden kann. Im Rahmen der Erfindung werden die so berechneten Parameter als DPR-Score bezeichnet, nach dem eine Normierung der Parameter erfolgt ist. Die
- 10 Normierung des DPR-Scores ist dabei so gewählt, dass ein DPR-Score von 0 einen an Diabetes erkrankten Diagnosehinweis liefert und ein DPR-Score von 1 durch die Probe eines gesunden Patienten generiert wurde. Findet eine Unterteilung der Zustände Diabetes, metabolisches Syndrom und Gesund in jeweils drei voneinander unabhängige Klassen statt, ergibt sich für die Klasse Diabetes ein DPR-Score im Mittelwert (10) von
- 15 0,25, für das metabolische Syndrom ein Mittelwert (10) von 0,65 und für die Klasse „gesund“ einen Mittelwert (10) von 0,87. In der dargestellten Box and Whiskers Plot zeigt dabei die Unterkante der Box jeweils den DPR-Score an unterhalb dieses Wertes 25 % der Parameter liegen. Der Mittelstrich in der Box wird als Median bezeichnet, wobei 50 % der Parameter kleiner als dieser Wert sind. Unterhalb des DPR-Scores der Ober-
- 20 kante der Box sind 75 % der Parameter zuweisbar. Des weiteren zeigen die jeweiligen Fehlerbalken in der Darstellung den Bereich, in dem 90 % der Parameter angeordnet sind. Wie den so dargestellten Daten zu entnehmen ist, sind in der Klasse Diabetes ca. 75 % der Parameter kleiner 0,25, sodass eine eindeutige Zuordnung dieser Proben in die Klasse möglich ist. Entsprechend liegen 50 % der Daten der Klasse metabolisches Syn-
- 25 drom unterhalb des DPR-Scores 0,87. 75 % der Daten, die der Klasse gesund angehören, liegen oberhalb des DPR-Scores 0,87. Wie der Darstellung beispielhaft zu entnehmen ist, liegen 50 % der Daten der Klasse metabolisches Syndrom in den Bereich $\text{DPR} > 0,87$, der ebenfalls einer Zuordnung zu der Klasse gesund zulässt. Die Klassen weisen folglich eine Schnittmenge an Daten auf, wobei eine Zuordnung eines Probenparameters zu
- 30 dieser Schnittmenge im Stand der Technik nicht möglich ist.

Erfindungsgemäß erfolgt eine Zuordnung einer unbekannten Probe jedoch zu einem Datensatz, der sich durch die Interpolation zwischen z. B. den Daten der Klasse „gesund“ und den Daten der Klasse „Diabetes“ berechnet. Im gezeigten Beispiel erfolgt die Interpolation der Daten durch eine Interpolation zwischen den DPR-Scores der

- jeweiligen Klassen. Erfolgt eine Interpolation zwischen den DPR-Scores der Klasse gesund und Diabetes, ergibt sich ein Kurvenverlauf, der vorteilhafterweise den Krankheitsverlauf der Krankheit Diabetes widerspiegelt. Hierbei zeigt sich, dass der DPR-Score der Klasse des metabolischen Syndroms zu dem interpolierten Datensatz zwischen den
- 5 Klassen gesund und Diabetes zuordnen lässt. Auf diese Weise wird ein Zusammenhang der drei verschiedenen Klassen deutlich, der auf einen medizinischen Zusammenhang der Krankheiten Rückschlüsse zulässt. Entsprechend der Auswertung ist folglich die Diagnose des metabolischen Syndroms in den Krankheitsverlauf der Krankheit Diabetes einzuordnen und ist somit als eine Vorstufe der Krankheit Diabetes anzusehen.
- 10 Figur 3 zeigt die Auftragung zweier Hauptkomponente eines multivariaten Auswerteverfahrens, die bei der Analyse einer Probe generiert wurden. Eine Auftragung von Hauptkomponenten eines multivariaten Auswerteverfahrens gegeneinander lässt die Kennzeichnung von Bereichen zu, die jeweils einer Klasse A oder B zugeordnet werden können. Hierbei liegen die viereckig dargestellten Datenpunkte im Bereich der Klasse B
- 15 sowie die kreisförmig dargestellten Datenpunkte den Bereich der Klasse A wiedergeben. Die Diskriminanzfunktion (20) grenzt die Bereiche der jeweiligen Klassen ab. In einem Bereich in unmittelbarer Nähe der Diskriminanzfunktion, kann eine Zuordnung zur Klasse A oder B nicht mehr eindeutig erfolgen. In diesem Bereich werden die Datenpunkte jeweils mit einem Kreuz gekennzeichnet. Diese Datenpunkte bilden eine
- 20 Schnittmenge zwischen den Klassen A und B, deren Zuordnung zu einer Klasse im Stand der Technik mittels bekannter Verfahren nicht möglich war.

Figur 4 stellt eine hieraus abgeleitete Quantifizierung des Krankheitszustandes dar, der sich aus einer Interpolation zwischen den DPR-Scores von Klasse A und B ergibt. Es erfolgt eine Auftragung der Anzahl der Proben, die vermessen wurde, gegen den berechneten DPR-Score. Hierbei zeigt sich, dass ein erheblicher Anteil Proben in dem Bereich der Schnittmenge liegt, sodass im Stand der Technik keine Zuordnung zu einer Klasse möglich ist. Mit Hilfe des erfindungsgemäßen Verfahrens ist eine Zuordnung dieser Proben in den abgebildeten Krankheitsverlauf problemlos möglich. Eine Probe mit einem DPR-Score von 0,6 kann somit der Diagnosehinweis gegeben werden, dass ein

25 Risiko, an der Krankheit der Klasse A zu erkranken, zu einem bestimmten Anteil gegeben ist. Dieses Risiko kann beispielsweise prozentual ausgedrückt, sodass ein DPR-Score von 0,6 eine 60%ige Wahrscheinlichkeit an der Klasse A zu erkranken wiedergibt. In dem aufgeführten Beispiel wird jeweils ein DPR-Score von 0 der Klasse B und von 1 der Klasse A zugeordnet. Alternativ kann auch der Abstand eines Datenpunktes zu der

30

Diskriminanzfunktion, wie in Figur 3 veranschaulicht, zur Bestimmung des Anteilswertes herangezogen werden. Hierbei wird der Abstand (21) als Lot zur Diskriminanzfunktion entsprechend dem vom DPR-Score aufgespannten Bereich von 0-1 zwischen Klasse A und B berechnet. Ebenso können in einer Graphik derartige Zusammenhänge veranschaulicht werden.

Figur 5 zeigt die Darstellung eines Krankheitsverlaufes, der im Stand der Technik (Circulation 80, 410 – 414 (1989)) als Braunwaldzyklus beschrieben wird. Der Braunwaldzyklus bildet die Verlaufsform einer koronaren Herzerkrankung ab und spiegelt somit Zusammenhänge zwischen verschiedenen Krankheitszuständen wider. Ausgehend von spezifischen Risikofaktoren oder anderen Frühstadien einer Krankheit äußert sich der Krankheitsverlauf zunächst an den Symptomen einer Arteriosklerose. Wird eine Arteriosklerose diagnostiziert kann dieser Krankheitszustand anfangs noch unbestimmt verlaufen. Zu diesem Zeitpunkt ist noch kein eindeutiger Zusammenhang zu einer schweren Herzerkrankung gegeben. Geht jedoch eine Arteriosklerose in eine koronare Herzerkrankung über, kann hierdurch bedingt eine Angina pectoris induziert werden. Gemäß dem Braunwaldzyklus kann dieser Krankheitszustand ein Herzinfarkt zur Folge haben.

Mit Hilfe des erfindungsgemäßen Verfahrens ist es möglich, Zusammenhänge zwischen Krankheiten, wie sie sich bereits im Braunwaldzyklus anhand von klinischen Untersuchungen darstellen lassen, direkt durch die Interpolation zwischen einem gesunden und einem kranken Zustand zu erfassen. Auf diese Weise können Diagnosehinweise einfach erstellt werden. Darüber hinaus können Zusammenhänge zwischen verschiedenen Krankheiten entsprechend einem Braunwaldzyklus erkannt werden. Beispielsweise werden somit Diagnosen von Krankheiten bereits in einem frühen Krankheitsstadium möglich. Die Früherkennung von Krebs oder TSE induzierten krankhaften Veränderungen wie BSE sind nur einige Beispiele, die die Anwendung des erfindungsgemäßen Verfahrens besonders vorteilhaft erscheinen lassen.

PATENTANSPRÜCHE

1. System, mit dem eine Zugehörigkeit einer Probe zu einer Klasse anteilmäßig bestimmt wird, sodass ein Diagnosehinweis zu einem Gesundheitszustand erstellt werden kann, beinhaltend
- 5
- eine Speichereinheit zur Speicherung der von einem Detektor erhaltenen Signale eines elektromagnetischen Spektrums, das durch Vermessen einer unbekannten Probe generiert wurde, und
 - eine Auswerteinheit zur Generierung eines Probenparameters mittels eines mul-
 - 10 tivariaten Auswerteverfahrens auf der Basis der Signale des elektromagnetischen Spektrums, und
 - zum Zuordnen des Probenparameter zu einem Datensatz, der sich aus der Interpolation zwischen mindestens einem Teil von ersten Daten und mindestens einem Teil von zweiten Daten berechnet, wobei
 - 15 - die ersten und zweiten Daten jeweils einer ersten und zweiten Klasse zugeordnet sind, und die ersten Daten einen ersten Parameter beinhalten, der anhand von elektromagnetischen Spektren von bekannten Proben, die der ersten Klasse angehören, mittels des multivariaten Auswerteverfahrens generiert wurde, sodass eine zuverlässige Zuordnung von Proben zu der ersten Klasse aufgrund des
 - 20 ersten Parameters möglich ist, und
 - die zweiten Daten einen zweiten Parameter beinhalten, der anhand von elektromagnetischen Spektren von bekannten Proben, die der zweiten Klasse angehören, mittels des multivariaten Auswerteverfahrens generiert wurde, sodass eine zuverlässige Zuordnung von Proben zu der zweiten Klasse aufgrund des zweiten
 - 25 Parameters möglich ist,
 - und die Auswerteinheit weiterhin zur Bestimmung eines Anteilswertes des Probenparameters an zumindest einem Teil des Datensatzes geeignet ist, wobei der Anteilswert eine Zugehörigkeit des Probenparameters zu einer Klasse wiedergibt, sodass ein Anteil der unbekannten Probe bestimmt wird, mit dem die
 - 30 unbekannte Probe einer Klasse angehört

- eine Ausgabeeinheit zur Anzeige der Zugehörigkeit einer Probe zu einer Klasse.
- 2. System gemäß Anspruch 1,
bei dem der erste und der zweite Parameter in einem Permanentpeicher der Speichereinheit gespeichert sind.
- 5 3. System gemäß Anspruch 1,
bei dem der Datensatz in einem Permanentpeicher der Speichereinheit gespeichert ist.
- 4. System nach einer der vorherigen Ansprüche,
bei dem Daten zu mehr als zwei Klassen gespeichert sind.
- 10 5. Analysesystem mit einem System nach einer der vorhergehenden Ansprüche,
das eine Strahlenquelle zum Bestrahlen einer Probe mit elektromagnetischer Strahlung beinhaltet.
- 6. Analysesystem nach Anspruch 5,
bei dem die Strahlenquelle elektromagnetische Strahlung im IR-Bereich emittiert.
- 15 7. Analysesystem mit einem System nach einer der vorhergehenden Ansprüche,
das einen Detektor zur Detektion der mit der Probe wechselwirkenden Strahlung beinhaltet.
- 8. Analysesystem nach Anspruch 7,
bei dem der Detektor Signale im Bereich von 1 μm - 25 μm detektiert.

9. Verfahren zur Auswertung eines elektromagnetischen Spektrums, wobei eine Zugehörigkeit des elektromagnetischen Spektrums zu einer Klasse anteilmäßig bestimmt wird, sodass ein Diagnosehinweis zu einem Gesundheitszustand erstellt werden kann, beinhaltend
- 5 - Ermittlung eines Probenparameters mittels eines multivariaten Auswerteverfahrens auf der Basis eines elektromagnetischen Spektrums, das durch eine Vermessung einer unbekannten Probe generiert wurde,
- 10 - Zuordnen des Probenparameters zu einem Datensatz, der sich aus der Interpolation zwischen mindestens einem Teil von ersten Daten und mindestens einem Teil von zweiten Daten berechnet, wobei die ersten und zweiten Daten jeweils einer ersten und zweiten Klasse zugeordnet sind,
- 15 - und die ersten Daten einen ersten Parameter beinhalten, der anhand von elektromagnetischen Spektren von bekannten Proben, die der ersten Klasse angehören, mittels des multivariaten Auswerteverfahrens generiert wurde, sodass eine zuverlässige Zuordnung von Proben zu der ersten Klasse aufgrund des ersten Parameters möglich ist,
- 20 - und die zweiten Daten einen zweiten Parameter beinhalten, der anhand von elektromagnetischen Spektren von bekannten Proben, die der zweiten Klasse angehören, mittels des multivariaten Auswerteverfahrens generiert wurde, sodass eine zuverlässige Zuordnung von Proben zu der zweiten Klasse aufgrund des zweiten Parameters möglich ist,
- 25 - Bestimmen eines Anteilswertes des Probenparameters an zumindest einem Teil des Datensatzes, wobei der Anteilswert eine Zugehörigkeit des Probenparameters zu einer Klasse wiedergibt, sodass ein Anteil der unbekannten Probe bestimmt wird, mit dem die unbekannte Probe einer Klasse angehört.
10. Verfahren gemäß Anspruch 9,
bei dem die erste Klasse den Gesundheitszustand eines Patienten mit Diabetes und die zweite Klasse den Gesundheitszustand eines Patienten ohne Diabetes charakterisiert und der Datensatz einen Probenparameter mit einschließt, der durch eine Probe eines Patienten mit metabolischem Syndrom generiert wird.
- 30

11. Verfahren gemäß Anspruch 9,
bei dem die Interpolation zwischen dem ersten und dem zweiten Parameter erfolgt.
12. Verfahren gemäß Anspruch 9,
bei dem der Anteil eines Probenparameters an einer Klasse prozentual berechnet
5 wird.
13. Verfahren gemäß Anspruch 9,
bei dem zwischen dem Datensatz und einem Krankheitsverlauf ein Zusammenhang
erstellt wird.
14. Verfahren gemäß Anspruch 9,
10 bei dem im Rahmen des multivariaten Auswerteverfahrens eine lineare Diskrimi-
nanzanalyse durchgeführt wird.
15. Verfahren gemäß Anspruch 9,
bei dem im Rahmen des multivariaten Auswerteverfahrens neuronale Netze
verwendet werden.
- 15 16. Verfahren gemäß Anspruch 9,
das Diagnosehinweise zur Früherkennung von Krebs- und TSE-induzierten krank-
haften Anomalien leistet.
17. System nach Anspruch 1,
das für ein Verfahren nach einem der Ansprüche 9- 16 geeignet ist.
- 20 18. Verfahren nach Anspruch 9,
bei dem ein System nach einem der Ansprüche 1-4 verwendet wird.

19. Datenträger mit einem Datensatz zur Analyse einer Probe beinhaltend

- einen Datensatz, der sich aus der Interpolation zwischen mindestens einem Teil von ersten Daten und mindestens einem Teil von zweiten Daten berechnet, wobei
- 5 - die ersten und zweiten Daten jeweils einer ersten und zweiten Klasse zugeordnet sind, und die ersten Daten einen ersten Parameter beinhalten, der anhand von elektromagnetischen Spektren von bekannten Proben, die der ersten Klasse angehören, mittels eines multivariaten Auswerteverfahrens generiert wurde, sodass eine zuverlässige Zuordnung von Proben zu der ersten Klasse aufgrund
- 10 des ersten Parameters möglich ist, und
- die zweiten Daten einen zweiten Parameter beinhalten, der anhand von elektromagnetischen Spektren von bekannten Proben, die der zweiten Klasse angehören, mittels des multivariaten Auswerteverfahrens generiert wurde, sodass eine zuverlässige Zuordnung von Proben zu der zweiten Klasse aufgrund des
- 15 zweiten Parameters möglich ist, wobei
- eine Zuordnung eines Probenparameters, der mittels des multivariaten Auswerteverfahrens auf der Basis eines elektromagnetischen Spektrums einer unbekannten Probe generiert wurde, zu dem Datensatz möglich ist.